

Some Dare Call It Conspiracy: Labeling Something a Conspiracy Theory Does Not Reduce Belief in It

Michael J. Wood

University of Winchester

“Conspiracy theory” is widely acknowledged to be a loaded term. Politicians use it to mock and dismiss allegations against them, while philosophers and political scientists warn that it could be used as a rhetorical weapon to pathologize dissent. In two empirical studies conducted on Amazon Mechanical Turk, I present an initial examination of whether this concern is justified. In Experiment 1, 150 participants judged a list of historical and speculative theories to be no less likely when they were labeled “conspiracy theories” than when they were labeled “ideas.” In Experiment 2 (N = 802), participants who read a news article about fictitious “corruption allegations” endorsed those allegations no more than participants who saw them labeled “conspiracy theories.” The lack of an effect of the conspiracy-theory label in both experiments was unexpected and may be due to a romanticized image of conspiracy theories in popular media or a dilution of the term to include mundane speculation regarding corruption and political intrigue.

KEY WORDS: conspiracy theories, labeling, stigma, belief

Since the mid-1990s, a growing psychological research tradition has generated a great deal of knowledge about the antecedents and consequences of beliefs in conspiracy theories. The term “conspiracy theory” itself, however, has received little explicit attention in the psychological literature, despite considerable interest from philosophers and political scientists in its precise meaning and implications (Bratich, 2002, 2008; Coady, 2006; deHaven-Smith, 2010, 2013; Husting & Orr, 2007). Calling something a conspiracy theory (or someone a conspiracy theorist) is seen as an act of rhetorical violence, a way of dismissing reasonable suspicion as irrational paranoia. For deHaven-Smith (2013), the conspiracy-theory label comes with such negative baggage that applying it has “the effect of dismissing conspiratorial suspicions out of hand with no discussion whatsoever” (p. 84). Husting and Orr (2007) likewise argued that applying the label “discredits specific explanations for social and historical events, regardless of the quality or quantity of evidence” (p. 131). Byford (2011) has lamented what he sees as a broadening of the meaning of “conspiracy theory” in recent decades: While the term once denoted speculation about secretive cabals controlling the course of world affairs, it has come to include “broader discourses of suspicion” (p. 151), such as routine mistrust of authority and concerns about the digital panopticon.

The intellectual stigma of conspiracy theorizing is not limited to academia. Wood and Douglas (2013) found that online commenters mostly used the term to characterize an opposing belief; commenters were strongly motivated to counterargue the label when others applied it to them, and very

few (even among those who believed that the U.S. government perpetrated the 9/11 attacks) were willing to apply the term to their own views. Politicians are clearly aware of the term's connotations: Both British Prime Minister David Cameron and New Jersey Governor Chris Christie have recently deflected accusations of impropriety by branding them as conspiracy theories (Benen, 2014; Helm & Boffey, 2011). It seems to be widely assumed, then, that labeling something a conspiracy theory makes it seem less believable—perhaps through association with a stereotyped view of conspiracy theories as paranoid and unfounded (Bratich, 2008).

There is good reason to believe that labels have power. Telling a mental health professional that a patient has been diagnosed with a particular disorder changes their evaluation of that patient's behavior—usually, not for the better (e.g., Fox & Stinnett, 1996). Likewise, preservice teachers are likely to judge a student's behavior much more harshly when told that the student has been given a diagnosis of Oppositional Defiant Disorder (e.g., Allday, Duhon, Blackburn-Ellis, & Van Dycke, 2010) or Attention Deficit Hyperactivity Disorder (e.g., Koonce et al., 2004). The effect of labeling on perception is evident at a basic cognitive level, too: in the well-known *prototype bias*, people recall previously seen stimuli as being more typical of their category than they actually are (e.g., Crawford, Huttenlocher, & Engebretson, 2000; Huttenlocher, Hedges, & Vevea, 2000). For instance, Huttenlocher, Hedges, and Duncan (1991) asked participants to categorize points in a circle, based on their positions, as belonging to one of four categories. The four categories essentially mapped onto the four quadrants of the circle. When later asked to remember the locations of the points they had categorized, participants showed a systematic bias: they recalled each point as being closer to the middle of its quadrant—in other words, as a more typical member of its category—than it actually was. The prototype bias influences higher-order judgments as well: People will distort representations of faces toward ethnic prototypes (Corneille, Huart, Becquart, & Brédart, 2004). This bias seems to be a plausible mechanism for the purported negative effects of the conspiracy-theory label. By most accounts, for the majority of people, the center of the conspiracy-theory category—the prototypical conspiracy theory—is seen as unfounded, paranoid, and silly (Bratich, 2008; deHaven-Smith, 2010, 2013; Husting & Orr, 2007). As such, perceptions of things categorized as conspiracy theories ought to drift toward that category prototype.

However, speculation about the cause of the negative effects of the conspiracy-theory label seems premature. Rather surprisingly, there has been no empirical investigation of whether the conspiracy-theory label actually has the impact people assume it does: No one has actually investigated whether calling something a conspiracy theory makes people believe it less. Moreover, some researchers have advised caution on this widely held assumption; while Uscinski and Parent (2014) find it likely that the net effect of the label is negative, they have pointed out that some people find the label “convincing and attractive” (p. 30) based on favorable and exciting portrayals of conspiracies in popular media. The importance of this issue is clear: If simply calling something a conspiracy theory really makes people take it less seriously, journalists should choose their words carefully—and the public should be on the lookout for propagandistic use of the term (deHaven-Smith, 2013).

To that end, the present study sought to determine whether the conspiracy-theory label is damaging to an idea's credibility, and if so, what some possible mechanisms might be. In the first of two experiments, I hypothesized that both speculative conspiracy theories and real historical conspiracies would seem less likely when called “conspiracy theories” than when called “ideas.”

Experiment 1

Method

Participants and Design. Ranging in age from 20 to 67 years ($M = 35.38$, $SD = 10.94$), 150 participants (61 female, 88 male, 1 transgendered) took part in Experiment 1. Recruitment was

accomplished via Amazon Mechanical Turk (MTurk), a web-based crowdsourcing service which allows participants to complete brief tasks for small payments. In this case, participants (all U.S. residents) were paid US\$1.00 each for taking part. MTurk provides fairly diverse and high-quality samples and has a good track record as an effective low-cost data source for psychological research (Paolacci & Chandler, 2014).

Experiment 1 followed an independent-groups experimental design, in which the independent variable was the type of label used, and the dependent variables were likelihood judgments of speculative and historical conspiracy theories. Participants were randomly assigned to one of two conditions: in the *conspiracy theory* condition ($N = 67$), they were asked to rate the likelihood of a variety of “conspiracy theories”; in the *idea* condition ($N = 83$), the same items were presented as “ideas” instead. The slightly unbalanced number of participants in each condition was the result of random assignment and resulted in a power of $1 - \beta = .86$ for a medium effect size (Cohen’s $d = .50$) under an independent-samples t -test. The overall sample size was determined by funding constraints, and all group difference analyses were preregistered.

Although 150 participants were paid via MTurk, 151 ultimately completed the questionnaire. The extra (unpaid) participant was excluded from analysis.

Materials. The questionnaire had two distinct components: a modified version of the Generic Conspiracist Beliefs scale (GCB; Brotherton, French, & Pickering, 2013) and five additional items regarding confirmed historical conspiracies. The order of presentation was counterbalanced within each condition, such that approximately half of all participants saw the GCB items first while the rest saw the historical items first.

Modified GCB. The GCB is a measure of conspiracist ideation. It makes no reference to specific conspiracy theories but comprises 15 nonspecific statements regarding abuses of power and cover-ups, for example, “The government permits or perpetrates acts of terrorism upon its own soil, disguising its involvement.” In the current study, each statement was converted to a question of likelihood, for example, “How likely is the [idea/conspiracy theory] that the government permits or perpetrates acts of terrorism upon its own soil, disguising its involvement?” Each question was answered by a 5-point Likert scale, ranging from 1 (not at all likely) to 5 (extremely likely).

The GCB has shown good convergent and discriminant validity in the past (Brotherton et al., 2013), and the modified form demonstrated high reliability in the present study (Cronbach’s $\alpha = .95$).

Historical conspiracies scale. The questionnaire contained an additional five items following the same format as the GCB. Rather than speculative conspiracy theories, these items referred to confirmed episodes in American history in which the government or other powerful actors engaged in conspiratorial conduct. As with the GCB items, these were referred to as either “ideas” or “conspiracy theories” depending on the condition; for instance, “How likely is the [idea/conspiracy theory] that the government has performed mind-control experiments on its own citizens without their consent?” refers to the MKULTRA program (see the appendix for the full scale). These items showed acceptable reliability (Cronbach’s $\alpha = .83$).

Procedure. The experiment was advertised on MTurk as “A short (~5min) survey about politics and history.” Prospective participants followed a link from MTurk to an external questionnaire page, hosted by QuestBack, where they were first presented with an informed-consent form. After agreeing to participate and providing demographic information, participants were presented with the GCB and historical conspiracy scales on the same page, in counterbalanced order, and asked to fill them out. Participants had to provide an answer for each item in order to complete the questionnaire. Upon completion, they were debriefed regarding the methods and hypothesis of the experiment and given a code to enter on the MTurk site in order to receive their payment.

Results

Participants rated the historical items as more likely ($M = 3.28$, $SD = 1.03$) than the GCB items ($M = 2.46$, $SD = .96$), 95% CI [.70, .93], Cohen's $d = .82$. However, the two measures showed a strong positive correlation, $r = .75$, 95% CI [.67, .81].

The mean likelihood rating of the GCB items was 2.46 ($SD = .99$) when they were labeled as "conspiracy theories" and 2.46 ($SD = .94$) when they were labeled as "ideas." This difference was not statistically significant, 95% CI [−.31, .31], Cohen's $d < .01$. Conspiratorial historical events were given a mean likelihood rating of 3.34 ($SD = 1.04$) when called conspiracy theories and 3.23 ($SD = 1.02$) when called ideas; this, too, was not a statistically significant difference, 95% CI [−.44, .23], Cohen's $d = -.10$.

Discussion

Experiment 1 did not find the hypothesized effects: Calling something a conspiracy theory failed to have any effect on people's evaluation of it. This was an unexpected result, and it runs counter to a long-standing assumption both within and outside academia.

However, it is possible that participants had simply already made their minds up about well-known historical conspiracies and the general topics covered in the GCB, and they were therefore unlikely to change their responses based solely on question wording. If this is so, the effect of calling something a conspiracy theory may only be detectable if the label is applied to something that the participant does not yet have a strong opinion on: Participants might have their own opinion on whether a particular proposition is best described as a conspiracy theory, one which is not overruled by the category label provided by the stimulus materials.

Moreover, it is likely that if there is a negative labeling effect, it will not take hold in everyone equally. People with a higher degree of conspiracist ideation tend to be more suspicious and mistrustful in general (Abalakina-Paap, Stephan, Craig, & Gregory, 1999), and conspiracy theories can come together to form a world view or belief system (Goertzel, 1994; Swami, Chamorro-Premuzic, & Furnham, 2010; Wood, Douglas, & Sutton, 2012)—indeed, this sort of conspiracist ideation is exactly what the GCB was designed to measure (Brotherton et al., 2013). Someone who believes in many different conspiracy theories might not be discouraged from adopting new beliefs by a "conspiracy theory" label, as that same label has already been applied to their own beliefs in other domains, and they still hold those beliefs regardless. On the other hand, it would be reasonable to expect that someone who is highly anticonspiracist would be motivated to avoid believing in anything to which the label is applied. In the language of the prototype bias, their category prototype for conspiracy theories would generally be less favorable than others', and therefore more likely to produce the sort of unfavorable category bias that could result in a negative labeling effect.

It is interesting to note that people's judgments about the truth of veridical historical events, such as the MKULTRA program, closely matched their judgments about the truth of the more speculative GCB items, such as a UFO cover-up. There are a few possible reasons for this; For instance, people who are aware of past malfeasance by powerful actors in society might extrapolate from known abuses of power to more speculative ones. Alternatively, people with more conspiracist world views might be more likely to seek out information on criminal acts carried out by officials in the past, while those with less conspiracist world views might ignore or reject such information. It is not clear to what degree the historical scale responses reflect actual knowledge of MKULTRA, COINTELPRO, etc., rather than world-view-based likelihood judgments. Undoubtedly some participants had not heard of the historical items and simply responded in line with their personal biases, but the generally higher ratings for the historical items suggest either generally higher plausibility or that at least some people knew that they referred to actual events.

To address the limitations of Experiment 1 in regard to the labeling effect, I conducted a second experiment about a novel subject, with a sample large enough to detect a small effect at 80% power. The primary hypothesis was the same as in Experiment 1: Calling something a conspiracy theory should result in people believing it less than if a more neutral term is used. A secondary hypothesis proposed that the effect of the conspiracy-theory label on belief is moderated by general conspiracist ideation, such that the negative effect of calling something a conspiracy theory is attenuated by the degree to which someone holds a conspiracist world view.

Experiment 2

Method

Participants and Design. Eight hundred and two participants (318 female, 484 male), ages 18–75 years ($M = 32.28$, $SD = 11.33$), took part in Experiment 2. All participants were U.S. residents recruited via MTurk and were paid US\$0.80 each in exchange for participating.

In an independent-groups experimental design, participants were randomly assigned to one of two conditions. In the experimental condition, 404 participants were asked about “conspiracy theories;” in the control condition, the remaining 398 participants were asked about “corruption allegations.” The dependent variable was a measure of composite endorsement similar to that used by Wood et al. (2012). Conspiracist ideation was also measured as a potential moderator.

Due to the cost associated with the large sample size, Experiment 2 made use of a sequential analysis technique (Lakens & Evers, 2014). Using an O’Brien-Fleming spending function (O’Brien & Fleming, 1979), a small portion of α was set aside so that an interim analysis could be carried out after collecting half of the final sample ($N = 401$, $\alpha_{0.5} = .00305$); if this analysis revealed a large enough difference between conditions, there would be no need to carry out any further data collection. The final sample size of $N = 802$ was based on an a priori power analysis with $\alpha_1 = .04695$ and a desired power of $1 - \beta = .80$ to detect a small effect size (Cohen’s $d = .20$).

Materials. The manipulation comprised a short mock news article about a fictitious political scandal in Canada. This particular subject matter was chosen to reduce the chance that participants in the U.S. sample would have a strong opinion on the veracity of the claims or be able to immediately recognize the news story as made-up. The body of the article was the same for all participants and described the ruling Conservative party’s denial of accusations that they had misappropriated public money to fund a recent reelection campaign. In the experimental condition, the headline was *Conspiracy Theories Emerge in Wake of Canadian Election Result*; in the control condition, the words *Conspiracy Theories* were replaced by *Corruption Allegations*. The article was presented in a realistic-looking news website template and was accompanied by an image of Prime Minister Stephen Harper.

The dependent measure, adapted from the composite endorsement scale used by Wood et al. (2012), was a series of six Likert-scaled items displayed on a separate screen from the mock article. On a scale from 1 (*not at all*) to 7 (*very much*), participants were asked to rate the degree to which they thought the accusations against the Canadian government were likely, plausible, convincing, worth considering, interesting, and coherent. Composite endorsement for each participant was defined as the mean of all of these items except interestingness (Wood et al., 2012) and showed good reliability (Cronbach’s $\alpha = .88$). The moderator, conspiracist ideation, was measured using the standard GCB scale (Brotherton et al., 2013) and also showed high reliability (Cronbach’s $\alpha = .93$).

Procedure. Experiment 2 was advertised on MTurk. The description informed participants that “[they would] be asked to read a brief news article, give [their] opinion on it, and answer a few questions about [their] views of the world. (~5 min).” As in Experiment 1, the MTurk page directed participants to an external questionnaire (hosted by Qualtrics), where they were

presented with an informed-consent form. After agreeing to participate and providing demographic information, participants were shown a screen telling them to read the following article carefully and not to seek information on the topic of the article elsewhere until the study was complete. They were then shown the article and had to answer a basic comprehension question about its content before moving on. The next screen contained the composite endorsement measure, and the screen after that administered the GCB. This ordering was the same across participants in order to prevent the content of the GCB questions from affecting endorsement scores. Finally, participants were shown a debriefing screen and given a code with which to claim their payment from MTurk.

Results

Interim group-differences analysis. As planned, data collection was initially halted at $N = 401$ to perform an interim analysis. Participants in the corruption-allegations condition showed a mean endorsement of 4.87 ($SD = 1.06$), while participants in the conspiracy-theories condition had a mean of 4.83 ($SD = 1.04$). This difference was not statistically significant, 99.695% CI $[-.17, .25]$ (recall $\alpha_{0.5} = .00305$), Cohen's $d = .04$, so data collection continued.

Final group-differences analysis. With the full sample ($N = 802$) collected, the corruption-allegations condition showed a mean composite endorsement of 4.89 ($SD = 1.07$), while the conspiracy-theory condition had a mean endorsement of 4.81 ($SD = 1.10$). This difference did not reach statistical significance, 95.305% CI $[-.07, .23]$ (recall $\alpha_1 = .04695$), Cohen's $d = .07$. The means and confidence intervals for each group are shown in Figure 1, along with the equivalent statistics for both measures in Experiment 1.

Moderation analysis. The moderation hypothesis was tested via multiple linear regression with composite endorsement as the dependent variable (Cohen, Cohen, West, & Aiken, 2002). Condition, GCB score (centered), and their product were entered as predictors. While higher GCB scores were significantly associated with endorsement of the allegations, $\beta = .17$, 95% CI $[.07, .27]$, the product term was not related to endorsement, $\beta = .03$, 95% CI $[-.07, .13]$. This indicates that the relationship between condition and endorsement was not moderated by GCB score.

Comprehension question analysis. Out of 802 participants, 156 got the comprehension question wrong. Every incorrect answer mixed up the two politicians referred to in the mock article. A separate analysis of only those participants who answered correctly still showed no difference in endorsement between the conspiracy-theories condition ($M = 4.74$, $SD = 1.10$) and the corruption-allegations condition ($M = 4.86$, $SD = 1.11$), 95.305% CI $[-.05, .29]$, Cohen's $d = .11$.

Discussion

As in Experiment 1, the hypothesis that the conspiracy-theory label would result in attenuated belief was not supported. Even when someone's very first exposure to an allegation of political corruption is seeing it branded as a conspiracy theory, they are no less likely to take it seriously than if it is instead called a corruption allegation. Moreover, the anticipated moderation effect failed to materialize: Even though people with more conspiracist world views generally took the allegations more seriously, the "conspiracy theory" label was just as powerless for conspiracy believers as it was for conspiracy skeptics. There was likewise no statistical difference between conditions when limiting analysis to only those people who answered the comprehension question correctly.

One potential limitation of Experiment 2 is that the alleged perpetrator in the stimulus article was the Canadian Conservative Party. Since MTurk participants are more liberal than the general population (Berinsky, Huber, & Lenz, 2012), this raises the possibility that a genuine difference in

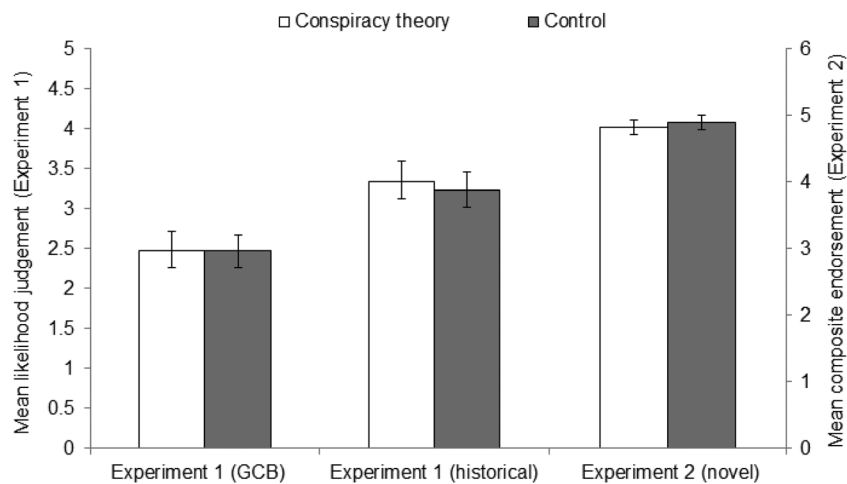


Figure 1. Mean likelihood judgements (Experiment 1) and composite endorsement (Experiment 2) for each labeling condition across both experiments. Error bars represent 95% confidence intervals.

endorsement between conditions may have been attenuated by an intergroup effect.¹ For instance, people may be motivated to ignore the negative connotations of conspiracy theories when conspiracy theorizing allows them to attribute malfeasance to outgroup members. When an outgroup member accuses ingroup members of dishonest conduct, on the other hand, people may be motivated to dismiss the allegation as an unfounded conspiracy theory. Thus, it is possible that the effect of conspiracy-theory labeling is moderated by group identification, such that the label only makes something less believable when it offers a way of dealing with group-esteem threat (Branscombe, Ellemers, Spears, & Doosje, 1999).

However, I argue that this does not plausibly account for the failure to detect an effect in Experiment 2. Although MTurk samples from the United States tend to skew liberal, they still contain a substantial percentage of conservative and Republican-identified participants (Berinsky et al., 2012). The power of Experiment 2 is high enough that an effect of practical significance within a reasonably sized subset of the sample should still have been detected. Moreover, the stimulus materials were specifically chosen to minimize American participants' investment in the subject, as Canadian politics are likely of minimal interest to most of the sample.

While participants were instructed not to look up information on the supposed scandal before completing the questionnaire, there was no means of enforcing this, so it is quite possible that some accuracy-motivated respondents sought out information elsewhere. This could have introduced some noise variance and reduced the power of the study. However, MTurk workers depend on rapid completion of many different tasks to earn a reasonable hourly wage (Berinsky et al., 2012), so the opposite motivation should have been more salient for most participants. In addition, the scandal was entirely made-up; any participants who looked for confirmation of the scandal online would have failed to find any, and as a result would presumably consider it less likely to be true. If a significant

¹ MTurk workers also have a higher rate of unemployment and a lower median income than the general population (for a review, see Paolacci & Chandler, 2014). It is possible that the stigma of the label is only absent among those who lack much power in society and that a labeling effect would be evident only among those with a relatively low level of anomie or a generally high socioeconomic status. The lack of a moderation effect for conspiracist ideation in Experiment 2 suggests that if such an effect exists, it is not due to the observed correlation between anomie and conspiracy belief (Abalakina-Paap et al., 1999); nevertheless, future work would undoubtedly benefit from more socioeconomically diverse sampling.

number of participants had looked for information elsewhere, the mean endorsement should therefore be quite low. However, mean endorsement lay well above the midpoint in both conditions.

General Discussion

Two experiments found no evidence of a negative effect of calling something a conspiracy theory. Experiment 1 showed no evidence that the label had any effect on endorsement of general conspiracist views or beliefs in real historical conspiracies, and Experiment 2 failed to find an effect with a fictitious political scandal previously unknown to participants and a large enough sample to detect a small effect with 80% power. Of course, it is possible that a weak effect of the label exists and was not detected or that the effect is contingent upon some circumstances not met by the study materials. Yet such an effect would be a small and slippery one, not reliably elicited even with a large sample and a blank slate of a political scandal to work on. The effect failed to show itself even among those who paid close attention to the stimulus materials, and there was no evidence that an overall negative view of conspiracy theories rendered the label any more effective in discouraging belief. Taken together, contrary to deHaven-Smith (2013) and Husting and Orr (2007), Experiments 1 and 2 suggest that the conspiracy-theory label possesses far less rhetorical power than previously assumed.

This is surprising, given past results showing that people actively counterargue the label when others apply it to their beliefs (Wood & Douglas, 2013). However, as noted by Uscinski and Parent (2014), the term has some positive connotations that may cancel out some of the intellectual stigma associated with it. Conspiracy theories are a common topic in popular media: Many films, video games, and television shows feature the protagonists fighting against shadowy conspiracies of frightening reach and power. Conspiracy theories have a kind of romanticism to them; they often take a utopian, Manichean view of the world, where humanity lives in a natural state of peace and harmony that was disrupted only by the conspirators' sinister meddling (Oliver & Wood, 2014; Willman, 2002). If the conspirators can be foiled by an awakened populace, a golden age will result. Despite their perceived lack of intellectual rigor, conspiracy theories are nevertheless attractive—perhaps attractive enough that the label prompts ambivalence rather than scorn when applied to an ambiguous event.

While conspiracy *theories* might not be viewed very negatively, it is possible that conspiracy *theorists* are viewed quite negatively. Unlike the conspiracies they believe in, conspiracy theorists do not often enjoy a positive or romanticized portrayal in popular media: their fictional counterparts are generally unbalanced, paranoid, dogmatic, and antisocial (Bratich, 2002, 2008). The label may therefore be effective in shaping opinions of people, if not of theories themselves. Husting and Orr (2007) have proposed that the conspiracy-theory label exerts its influence by shifting attention away from the validity of the claims and toward the competence and credibility of the person making them. However, the stimulus materials in Experiment 2 identified parties who took the allegations seriously and demanded an investigation; if these parties were tarred by their association with the “conspiracy theory,” that negative evaluation nevertheless did not spread back to the theory itself. It is possible, though, that participants' motivation to avoid being seen as conspiracy theorists themselves is not salient when responses are anonymous, as they were in these experiments. In a social situation in which they are evaluated by their peers, people might be more likely to reject something labeled a conspiracy theory for fear of negative social consequences.

Another potential explanation for these null effects concerns the mechanism by which category labels influence perception. People make categorization decisions independently—they do not always rely on others to tell them whether something is a conspiracy theory or not. In some cases, as with the GCB items regarding UFO cover-ups, people might have applied the label regardless of the manipulation. Conversely, allegations of straightforward political corruption like those in Experiment 2 might prompt counterargument against the label: This isn't really a conspiracy theory; it's just politics as

usual. Seeing the label applied by a third party might only be effective in more ambiguous cases, where the thing in question is neither obviously a conspiracy theory nor too mundane to merit the label under normal circumstances. Moreover, the prototype bias does not seem to affect initial perceptions of an object—rather, it arises when a stimulus is reconstructed from memory, as the reconstruction process incorporates information from the category label (Crawford et al., 2000). Likewise, Darley and Gross (1983) proposed that the effect of labels on person perception depends on biased assimilation of evidence: Applying a label has little initial effect but shapes subsequent perceptions in a way that makes the label seem justified in retrospect. Rather than having an immediate effect, the conspiracy-theory label might color perceptions of evidence that arises at a later time. However, the mock article in Experiment 2 provided a reasonable amount of ambiguous information for participants to read after the initial label was given in the article’s headline. This additional information could easily have been assimilated in a biased manner according to which label participants were assigned, yet there was still no evidence of an effect. Moreover, in Experiment 2, participants were asked for their endorsement on a separate screen from the stimulus article, meaning that some reconstruction from memory would have been necessary in order to make the required judgments. Indeed, simply showing stimuli on a separate screen from the recall task was enough to produce a robust prototype bias in Crawford et al. (2000). Regardless, it would be revealing (if methodologically tricky) to investigate whether labeling leads to biased information processing over a longer period of time and a greater amount of ambiguous information.

Finally, it is possible that the conspiracy-theory label has simply lost some of the power that it once had. While the effectiveness of the manipulation may not have been moderated by beliefs in conspiracy theories, it is quite possible that even people who are skeptical of the sorts of conspiracy theories mentioned in the GCB are sympathetic to the idea of conspiracies in general: Perhaps the conspiracy-theory label’s common meaning extends beyond the subject matter of the GCB to include general speculation about political intrigue. Byford’s (2011) concern about the dilution of the term may be well-founded: as the cases of David Cameron (Helm & Boffey, 2011) and Chris Christie (Benen, 2014) demonstrate, the label is sometimes used defensively by politicians to associate relatively mundane suspicions with fanciful speculation about world-controlling cabals. Rather than putting those suspicions to rest, this tactic may instead have caused a reevaluation of the label itself: For many, it may have prompted the question of whether conspiracy theories might be on the right track after all.

Preregistration

Hypotheses and primary analyses for these experiments were preregistered via the Open Science Framework website. Time-stamped hypotheses, analysis plans, materials, anonymized SPSS raw data, and descriptions of the experimental procedure are available at <https://osf.io/y8iu3/> (Experiment 1) and <https://osf.io/tqxp/> (Experiment 2).

ACKNOWLEDGMENTS

Many thanks to Lee Basham and two anonymous reviewers for invaluable feedback on an earlier draft of this article. Correspondence concerning this article should be addressed to Michael Wood, University of Winchester, Sparkford Road, Winchester SO22 4NR, UK. E-mail: Michael.Wood@winchester.ac.uk

REFERENCES

- Abalakina-Paap, M., Stephan, W. G., Craig, T., & Gregory, W. L. (1999). Beliefs in conspiracies. *Political Psychology*, 20, 637–647. doi:10.1111/0162-895X.00160

- Allday, R. A., Duhon, G. J., Blackburn-Ellis, S., & Van Dycke, J. L. (2011). The biasing effects of labels on direct observation by preservice teachers. *Teacher Education and Special Education, 34*, 52–58. doi:10.1177/0888406410380422
- Benen, S. (2014, January 15). Reflecting conspiracy theories. *The Rachel Maddow Show*. Retrieved from www.msnbc.com/rachel-maddow-show/reflecting-conspiracy-theories
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*, 351–368. doi:10.1093/pan/mpr057
- Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers & R. Spears (Eds.), *Social identity: Contexts, commitment, content* (pp. 35–59). Oxford, UK: Blackwell Science.
- Bratich, J. Z. (2002). Injections and truth serums: AIDS conspiracy theories and the politics of articulation. In Knight, P. (Ed.), *Conspiracy nation: The politics of paranoia in postwar America* (pp. 133–156). New York, NY: New York University Press. doi:10.2307/3089853
- Bratich, J. Z. (2008). *Conspiracy panics: Political rationality and popular culture*. New York, NY: State University of New York Press.
- Brotherton, R., French, C., & Pickering, A. D. (2013). Measuring belief in conspiracy theories: The Generic Conspiracist Beliefs scale. *Frontiers in Psychology, 4*, 279. doi: 10.3389/fpsyg.2013.00279
- Byford, J. (2011). *Conspiracy theories: A critical introduction*. Basingstoke, UK: Palgrave Macmillan. doi:10.1057/9780230349216
- Coady, D. (2006). *Conspiracy theories: The philosophical debate*. Aldershot, UK: Ashgate. doi:10.1057/9780230349216.0005
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum. doi:10.4324/9780203774441
- Corneille, O., Huart, J., Becquart, E., & Brédart, S. (2004). When memory shifts toward more typical category exemplars: Accentuation effects in the recollection of ethnically ambiguous faces. *Journal of Personality and Social Psychology, 86*, 236–250.
- Crawford, L. E., Huttenlocher, J., & Engebretson, P. H. (2000). Category effects on estimates of stimuli: Perception or reconstruction? *Psychological Science, 11*, 280–281. doi:10.1111/1467-9280.00256
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*, 20–33. doi:10.1037/0022-3514.44.1.20
- deHaven-Smith, L. (2010). Beyond conspiracy theory: Patterns of high crime in American government. *American Behavioral Scientist, 53*, 795–825. doi:10.1177/0002764209353274
- deHaven-Smith, L. (2013). *Conspiracy theory in America*. Austin: University of Texas Press.
- Fox, J. D., & Stinnett, T. A. (1996). The effects of labeling bias on prognostic outlook for children as a function of diagnostic label and profession. *Psychology in the Schools, 33*, 143–152. doi:10.1002/(SICI)1520-6807(199604)33:2<143::AID-PITS7>3.0.CO;2-S
- Goertzel, T. (1994). Belief in conspiracy theories. *Political Psychology, 15*, 731–742. doi:10.2307/3791630
- Helm, T., & Boffey, D. (2011, July 24). How phone hacking took the shine off the Prime Minister. *The Guardian*. Retrieved from <http://www.theguardian.com/media/2011/jul/24/david-cameron-hacking-conservative-fallout>
- Husting, G., & Orr, M. (2007). Dangerous machinery: “Conspiracy theorist” as a transpersonal strategy of exclusion. *Symbolic Interaction, 30*, 127–150. doi:10.1525/si.2007.30.2.127
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98*, 352–376.
- Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129*, 220–241.
- Koonce, D. A., Cruce, M. K., Aldridge, J. O., Langford, C. A., Sporer, A. K., & Stinnett, T. A. (2004). The ADHD label, analogue methodology, and participants' geographic location on judgments of social and attentional skills. *Psychology in the Schools, 41*, 221–234. doi:10.1002/pits.10150
- Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science, 9*, 278–292. doi: 10.1177/1745691614528520
- O'Brien, P. C., & Fleming, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics, 35*, 549–556. doi: 10.2307/2530245
- Oliver, J. E., & Wood, T. J. (2014). Conspiracy theories and the paranoid style(s) of mass opinion. *American Journal of Political Science, 58*(4), 952–966. doi:10.1111/ajps.12084
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science, 23*, 184–188. doi:10.1177/0963721414531598

- Swami, V., Chamorro-Premuzic, T., & Furnham, A. (2010). Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Applied Cognitive Psychology*, 24, 749–761. doi:10.1002/acp.1583
- Uscinski, J. E., & Parent, J. M. (2014). *American conspiracy theories*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199351800.001.0001
- Willman, S. (2002). Spinning paranoia: The ideologies of conspiracy and contingency in postmodern culture. In Knight, P. (Ed.), *Conspiracy nation: The politics of paranoia in postwar America* (pp. 21–39). New York, NY: New York University Press.
- Wood, M. J., & Douglas, K. M. (2013). “What about building 7?” A social psychological study of online discussion of 9/11 conspiracy theories. *Frontiers in Psychology*, 4, 409. doi:10.3389/fpsyg.2013.00409
- Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and alive: Beliefs in contradictory conspiracy theories. *Social Psychological and Personality Science*, 3, 767–773. doi:10.1177/1948550611434786

Appendix

Historical conspiracy items from Experiment 1.

1. How likely is the [idea/conspiracy theory] that the government has performed mind-control experiments on its own citizens without their consent?
2. How likely is the [idea/conspiracy theory] that government agencies have recruited journalists into a secret propaganda network in order to influence the media?
3. How likely is the [idea/conspiracy theory] that government agencies have illegally used surveillance, infiltration, and wrongful imprisonment in order to discredit domestic political groups that they considered potentially threatening?
4. How likely is the [idea/conspiracy theory] that the IRS and other agencies have been used by presidential administrations to harass political rivals?
5. How likely is the [idea/conspiracy theory] that federal government officials have taken bribes from industry leaders in exchange for favorable government contracts?

Item 1 refers to the MKULTRA experiments; Item 2 refers to Operation Mockingbird; Item 3 refers to COINTELPRO; Item 4 refers to several incidents through U.S. history, including the use of the IRS by President Nixon; Item 5 likewise refers to a variety of incidents, including the Teapot Dome scandal.